Signal Processing 91 (2011) 1066–1070

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Fast communication

Adaptive sigmoidal plant identification using reduced sensitivity recursive least squares ☆

A. Navia-Vázquez *, R. Díaz-Morales

Department of Signal Processing and Communications, Univ. Carlos III Madrid, Avda Universidad 30, 28911-Leganés, Madrid, Spain

ARTICLE INFO

Article history:
Received 10 February 2010
Received in revised form
27 May 2010
Accepted 21 September 2010
Available online 29 September 2010

Keywords:
Adaptive
Nonlinear
Logistic
Sigmoid
Plant
System
Identification
Weighted Recursive Least Squares

ABSTRACT

Logistic models, comprising a linear filter followed by a nonlinear memoryless sigmoidal function, are often found in practice in many fields, e.g., biology, probability modelling, risk prediction, forecasting, signal processing, electronics and communications, etc., and in many situations a real time response is needed. The online algorithms used to update the filter coefficients usually rely on gradient descent (e.g., nonlinear counterparts of the Least Mean Squares algorithm). Other algorithms, such as Recursive Least Squares, although promising improved characteristics, cannot be directly used due to the nonlinearity in the model. We propose here a modified Recursive Least Squares algorithm that provides better performance than competing state of the art methods in an adaptive sigmoidal plant identification scenario.

1. Introduction

Structures comprising a linear filter followed by a nonlinear memoryless sigmoidal function are useful models in many application fields (biology, probability modelling, risk prediction, forecasting, signal processing, electronics and communications). In some cases, the input data does not follow a stationary distribution and adaptive algorithms are needed to adjust the model weights. The most common family of such algorithms are those based on gradient descent that follow directions in the weight space opposite to the gradient of the error surface to find a suitable solution for the weights (e.g., nonlinear counterparts of the Least Mean Squares (LMS) algorithm). However, such methods are slow to converge and very sensitive to highly correlated inputs. Other families of algorithms, such as those based on minimizing a Least Squares cost function (e.g., Recursive Least Squares, RLS), although showing improved characteristics, cannot directly be used due to the nonlinearity in the model. Some solutions have been proposed using piecewise approximations of the sigmoid using Taylor's expansions, and hence they are suboptimal [1,3–5,7,8]. An improved approach, named as Non-Linear RLS (NL-RLS), has been proposed in [6] and since it does not rely on any approximation it outperforms the aforementioned methods. We propose here a modified Recursive Least Squares algorithm that provides better performance than competing state of the art methods in an adaptive sigmoidal plant identification scenario, as will be shown in the experimental section.

2. The proposed Reduced-Sensitivity RLS algorithm

The task of plant identification with sigmoidal function at the output of the filter has been depicted in Fig. 1(a). The top branch represents the plant output generation given an input

☆ This work has been supported by Spain Government (Grant TEC2008-02473).

* Corresponding author. Tel.: +34 1 624 59 77; fax: +34 1 624 87 49.
E-mail address: navia@tsc.uc3m.es (A. Navia-Vázquez).

0165-1684/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2010.09.017

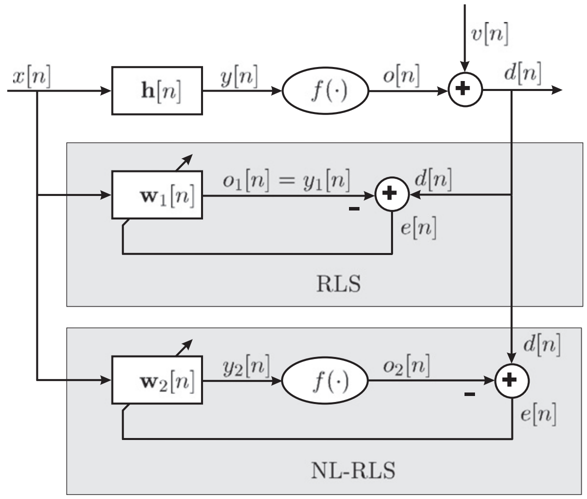

Fig. 1. Nonlinear plant identification scenario. The plant identification schemes using RLS and NL-RLS algorithms are shown (grey boxes).

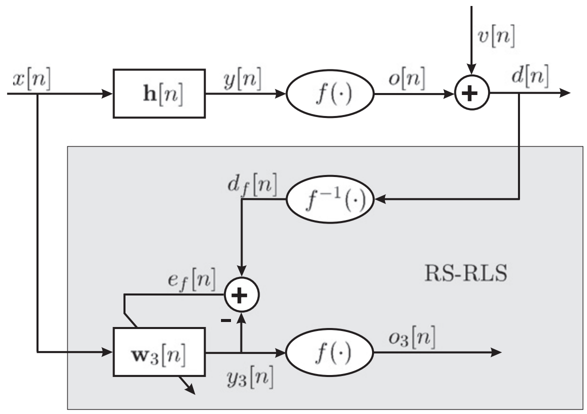

Fig. 2. The proposed RS-RLS algorithm for nonlinear plant identification.

signal $x[n]$ and a nonstationary impulse response $\mathbf{h}[n]$. The signal at the output of the filter is $y[n]=\mathbf{h}^T[n]\mathbf{x}[n]$, where $\mathbf{x}[n]=[x(n-N+1),\dots,x(n)]^T$ (N is the number of weights in the filter), the output of the sigmoidal function is $o[n]=f(y[n])=(e^{y[n]}-e^{-y[n]})/(e^{y[n]}+e^{-y[n]})$, and $o[n]$ is corrupted by additive white Gaussian noise $v[n]$ to finally produce $d[n]$, the desired plant output to be estimated. The grey box in Fig. 1 marked as "RLS" represents the standard plant identification approach using a linear filter adjusted using the RLS algorithm.

Such direct approach is bound to be suboptimal since it is designed for linear plant identification. The second grey box in Fig. 1 marked as "NL-RLS" represents the nonlinear solution proposed in [6], which claims to solve the problem without any linear approximation of the sigmoidal function. The filtering scheme proposed in this paper has been depicted in Fig. 2 (grey box marked as "RS-RLS") and, in what follows, we derive the weight update mechanism.

The error signal at the output of the nonlinearity in the RS-RLS case at time "n" is $e[n] = d[n]-o_3[n]=d[n]-f(\mathbf{w}_3^T[n]\mathbf{x}[n])$, which is nonlinear with respect to the weights. We propose to minimize the squared error before the nonlinearity, such that a closed form solution for the weights is available. Therefore, the desired output before the nonlinearity in the RS-RLS case at time "n" is $d_f[n]=f^{-1}(d[n])=1/2\ \ln((1+\hat{d}[n])/(1-\hat{d}[n]))$, where $\hat{d}[n]$ is $d[n]$ scaled to the range $(-1+\varepsilon,1-\varepsilon)$. The output of the filter also at time "n" is $y_3[n]=\mathbf{w}_3^T[n]\mathbf{x}[n]$ and the error before the nonlinearity is $e_f[n]=d_f[n]-y_3[n]$. However, we are actually interested in minimizing the error after the nonlinearity, and $e_f[n]$ is transformed by the sigmoid depending on the value of the filter output $y_3[n]$ (assuming that errors are much smaller than the filter output itself, which is the common situation). When $y_3[n]$ is close to zero, we are operating on the linear part of the sigmoid and error $e_f[n]$ appears unaltered after the nonlinearity. However, when $|y_3[n]|$ is large, the sigmoid attenuates the error $e_f[n]$. In general, the attenuation factor is proportional to the derivative of the sigmoid. Therefore a more realistic error measurement at the output of the sigmoid is

$$e[n]=(d_f[n]-y_3[n])f'(y_3[n]) \tag{1}$$

where $f'(y_3[n])$ is the derivative of the sigmoid function evaluated at $y_3[n]$, which can be efficiently computed as $f'(y_3[n])=1-(f(y_3[n]))^2$. We will incorporate this error definition in the formulation of the RS-RLS algorithm, as follows. The functional to be minimized at time n can be written in a Weighted Recursive Least Squares form:

$$C(\mathbf{w}_3[n])=\sum_{i=1}^{n}\lambda^{n-i}a(i)(d_f[i]-\mathbf{w}_3^T[n]\mathbf{x}[i])^2 \tag{2}$$

where $a(i)=(f'(y_3[n])^2$. The cost function in (2) is nonlinear (non-quadratic) with respect to the weights, since $a(i)$ depends on $y_3[n]$ and $y_3[n]$ is computed using the weights $\mathbf{w}_3[n]$, so it cannot be directly solved in explicit form. However, we may adopt an iterated weighted scheme consisting in assuming that $a(i)$ values do not depend on the weights, minimizing (2) to find new weights, updating the $a(i)$ values with the new weights, and iteratively repeating until convergence. This type of approaches are known in the literature as Iterated Weighted Least Squares (IWLS), and they are stable and converge reasonably fast [2,9].

Therefore, to solve (2) under the IWLS approach we compute the derivative of $C(\mathbf{w}_3[n])$ with respect to $\mathbf{w}_3[n]$, taking $a(i)$ as constants, and we set it equal to zero, to obtain:

$$\sum_{i=1}^{n}\lambda^{n-i}a(i)\mathbf{x}[i](d_f[i]-\mathbf{x}^T[i]\mathbf{w}_3[n])=0 \tag{3}$$

Eq. (3) leads to the well known linear system $\mathbf{R}_x[n]\mathbf{w}_3[n]=\mathbf{r}_{xd}[n]$ where

$$\mathbf{R}_x[n]=\sum_{i=1}^{n}\lambda^{n-i}a(i)\mathbf{x}[i]\mathbf{x}^T[i] \tag{4}$$

$$\mathbf{r}_{xd}[n]=\sum_{i=1}^{n}\lambda^{n-i}a(i)\mathbf{x}[i]d_f[i] \tag{5}$$

To recursively solve this system, we start from the direct recursion to estimate $\mathbf{R}_x[n]$ and $\mathbf{r}_{xd}[n]$, formulated